

# Protein Structure and Sequence Generation with Equivariant Denoising Diffusion Probabilistic Models

Year: 2022 | Citations: 198 | Authors: N. Anand, Tudor Achim

---

## Abstract

Proteins are macromolecules that mediate a significant fraction of the cellular processes that underlie life. An important task in bioengineering is designing proteins with specific 3D structures and chemical properties which enable targeted functions. To this end, we introduce a generative model of both protein structure and sequence that can operate at significantly larger scales than previous molecular generative modeling approaches. The model is learned entirely from experimental data and conditions its generation on a compact specification of protein topology to produce a full-atom backbone configuration as well as sequence and side-chain predictions. We demonstrate the quality of the model via qualitative and quantitative analysis of its samples. Videos of sampling trajectories are available at <https://nanand2.github.io/proteins> .